

# Polarisation of physics on global courses

**Allan L. Alinea**

Department of Physics, Osaka University, Japan, 560 0043

E-mail: [alinea@het.phys.sci.osaka-u.ac.jp](mailto:alinea@het.phys.sci.osaka-u.ac.jp)

**Wade Naylor**

International College & Department of Physics, Osaka University, Japan, 560 0043

E-mail: [naylor@phys.sci.osaka-u.ac.jp](mailto:naylor@phys.sci.osaka-u.ac.jp)

**Abstract.** Since October 2010, the Chemistry-Biology Combined Major Program (CBCMP), an international course taught in English at Osaka University, has been teaching small classes (no more than 20 in size). We present data from the Force Concept Inventory (FCI) given to first year classical mechanics students ( $N = 47$  students over three years) pre and post score, for a class that predominantly uses interactive engagement (IE), such as MasteringPhysics. Our findings show a  $G$ -factor improved score of about  $\sim 0.18$ , which is marginally about the average of a traditional based course. Furthermore, we analyse in detail a set of six questions from the FCI, involving the identification of forces acting on a body. We find that student answers tend to cluster about “polarising choices”—a pair of choices containing the correct choice and a wrong choice with the latter corresponding to a superset of forces in the former. Our results are suggestive that students have a good idea of the right set of forces acting on a given system but the inclusion of extra force(s) brings about confusion; something that may be explained by misleading ontological categorisation of forces. In an appendix we also comment on possible correlations between the pre/post score and the level of English ability on entry to the course.

*Keywords:* Physics Education; Interactive Engagement; Global Courses.

## 1. Introduction

In a typical calculus-based introductory mechanics course, a significant time is allotted in applying Newton's laws of motion. It follows that the skill in making Free Body Diagrams (FBD) (whether on paper or just inside one's head) is repetitively invoked throughout much of the course in solving mathematical and conceptual problems, and in trying to have a deeper understanding of the dynamics of several systems. Making a FBD is thus, one of the foundational skills a student is expected to acquire. Part of this foundational skill is the ability to identify the forces acting on a given body. If this very basic ability is not properly established, the student will have difficulty in gaining a deep understanding of Newtonian mechanics. For this reason, the need to look into this very simple matter of identifying forces cannot be overemphasised.

In the past couple of decades, much of the research involving students' understanding or the misconception of force and other related concepts in Newtonian mechanics focused on the "cure" in the form of interactive engagement (see for instance, [1], [2], [3], [4], [5]). Although this is an equally important area of research, the subject itself of student misconception regarding force is far from being a closed-case. As we shall see, there are still avenues (eg., the existence of "polarising choices" (subsection 3.2)) that need to be explored. New insights can be gained from this exploration through the administration of a concept inventory. In this work, we revisit the subject matter of understanding force with a special focus on the identification of forces through the use of the Force Concept Inventory (FCI).

The FCI developed by Hestenes, Wells, and Swackhamer in the early 1990s is an instrument to assess student's understanding or misconception of force [6], [7]. It consists of 30 multiple-choice questions each with five choices. After more than two decades of existence, there must be good reasons why it is still one of the most widely used concept inventories. Contained in the FCI are questions that suit our need to look into the subject of identifying force. This allows us to set aside the need to develop a new set of questions and focus on the analysis of the result of the FCI. The subjects of our testing are the students of the Chemistry-Biology Combined Major Program (CBCMP).

## 2. Methodology

The CBCMP is a program offered by Osaka University (Japan) that is geared for students coming from outside Japan. The student composition is multiracial in nature with approximately 50% coming from China (including on and off the mainland), 25% from Southeast Asia, and 25% from other countries including Japan. Classes are conducted in English with the instructor meeting the class once a week for a period of 1.5 hours and a total of 20 hours in class (this is to be contrasted with a standard physics course which typically has 30/40 hours).

Having a limited number of class hours per week, traditional lectures can be both ineffective and impractical, so to deal with this shortage of hours we have employed

Group	2011 ( $n = 11$ )	2012 ( $n = 18$ )	2013 ( $n = 18$ )	Total ( $N = 47$ )
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Pretest	17.1 (6.6)	16.6 (7.4)	14.7 (6.7)	16.0 (6.9)
Posttest	20.2 (7.0)	19.7 (6.3)	16.4 (7.0)	18.6 (6.8)
Gain*	0.24	0.23	0.12	0.18

**Table 1:** Result of the FCI for the three batches of entrants from 2011-2013. \*For the Normalised Gain standard deviation (SD) is not applicable.

MasteringPhysics [8]. In a typical class the students are given prelecture assignments (10 questions each) to read in advance before going to class and at every meet, the class starts with a set of prelecture conceptual questions some of which are discussed by the instructor. A short lecture then follows highlighting the main points of the scheduled subject of the meeting. For the remaining 60% of the class hours, students are paired to discuss and answer the assigned problems (conceptual or quantitative) from MasteringPhysics. While answering the problems, the instructor and teaching assistant attend to student concerns about the problems they are solving. They are given one week to finish the assigned set of problems where students on average spend about one to two hours doing their homework.

We administered the FCI at the beginning and at the end of the course on calculus-based introductory mechanics. The data corresponding to a total of 47 students, were taken over three years from October 2011 until February 2014. The three-year data were subjected to statistical analysis mainly involving  $t$ -test, graphical analysis, and calculation of the standard statistical parameters (eg., mean, standard deviation, etc).

### 3. Results and Discussion

#### 3.1. Result of the FCI as a Whole

Table 1‡ shows the result of the FCI for three batches of entrants to the CBCMP program. We would like to stress that given the number of students we have we found an overall difference in the means of the pre and post test scores in favor of the latter. To further verify this we used  $R$  (statistical software) [9] to find the paired student  $t$ -test result. At the 95% confidence level we found a  $p$ -value of  $4.1 \times 10^{-5}$  for 46 degrees of freedom with a mean difference of 8.6%. The null hypothesis was not satisfied and suggests the results do lead to a positive increase in the post scores.

The  $t$ -test only indicates whether there is a significant difference between the pre and post test scores. A better measure of student performance before and after instruction is provided by the *normalised gain*,  $G$  [2]. It is the difference between the pre and post test scores divided by the maximum possible score in the posttest relative

‡ For a recent article on which kind of concept test to implement and how to interpret the data see [10].

to the score in the pretest. Symbolically,

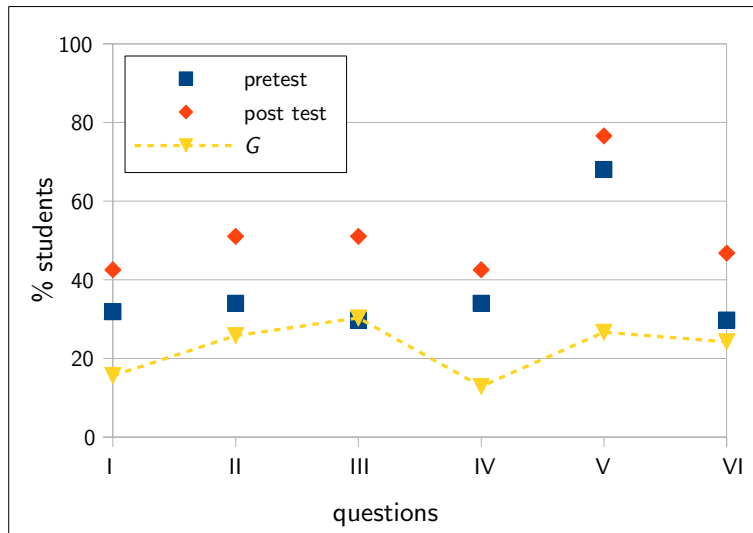
$$G = \frac{\langle \%S_f \rangle - \langle \%S_i \rangle}{100 - \langle \%S_i \rangle}, \quad (1)$$

where  $\%S_f$  and  $\%S_i$  are the final and initial scores respectively. **Table 1** shows the normalised gain for the three batches of students. On average, we find a result of  $G = 0.18$ . At first we found this result to be a little surprising, as we expected a higher gain in line with interactive engagement (IE) based courses [2], [11], [12]. In the appendix, we look into the possible correlation of the normalised gain with the English language ability.

### 3.2. Identifying Forces

The relatively low normalised gain may be attributed to a variety of factors. The complexity of the FCI however, does not give us much freedom to investigate a lot of things. The FCI centring around the idea of force connects to so many other structures of human understanding regarding force that we are bound to focus only on a small manageable part. Beyond that, we can say a lot of things from logic alone, but scarcity of experimental results on our part would simply hinder us in drawing a sound conclusion.

What we do in this work as already mentioned, is focus on the identification of forces. For this subject, we choose only segments of the FCI (labeled here as questions I to VI)<sup>§</sup> that fall under the mentioned category. The hope is to see an obvious pattern on how students identify force.

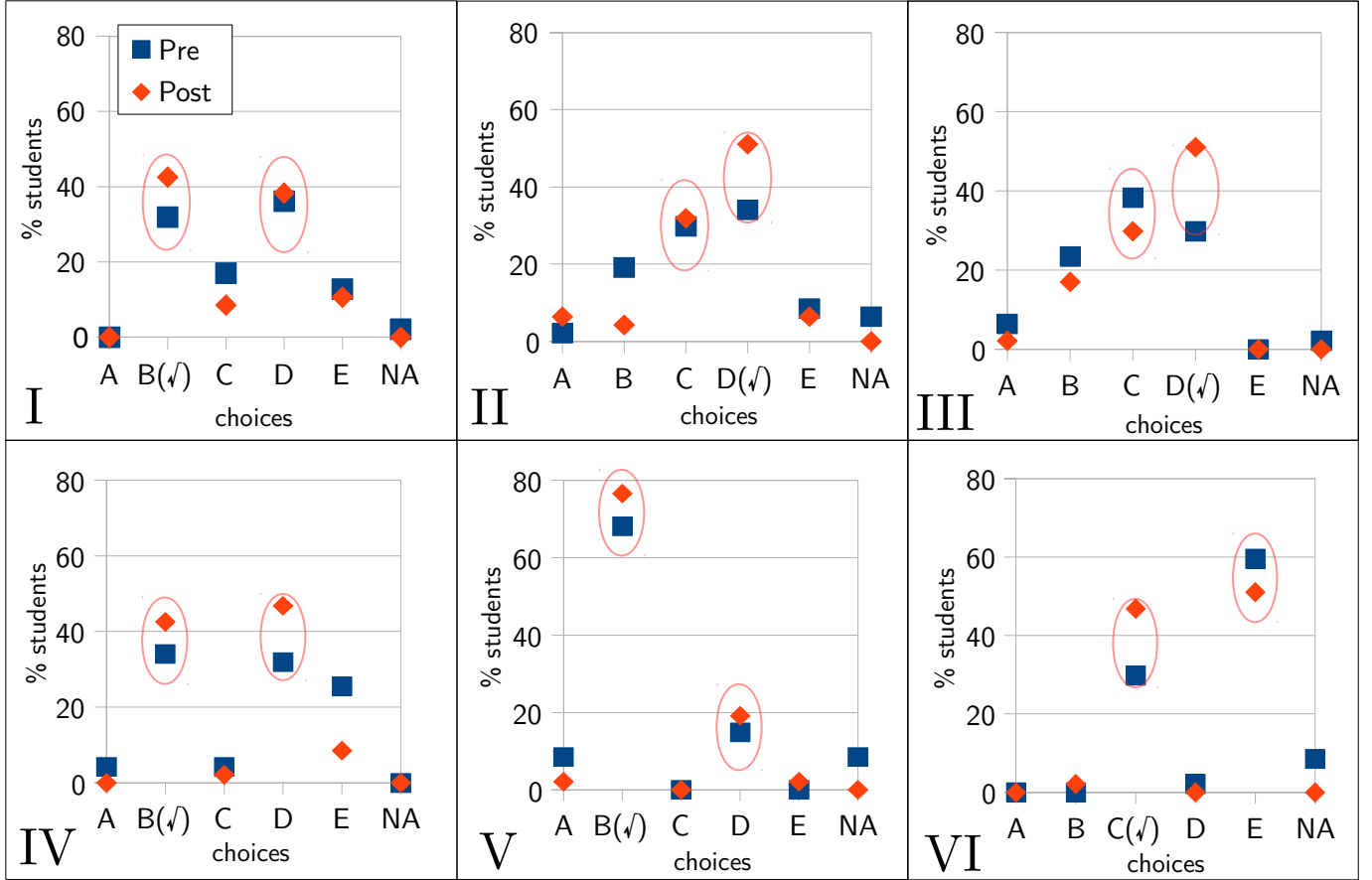


**Figure 1:** Normalised gain ( $G$ ) (expressed in percent) for selected questions involving the identification of forces acting on a given body.

Let us now look into the pre and post test scores. **Figure 1** shows the percent of students who got the correct answer for the six questions labelled I to VI involving

<sup>§</sup> These correspond to questions 5, 11, 13, 18, 29, and 30.

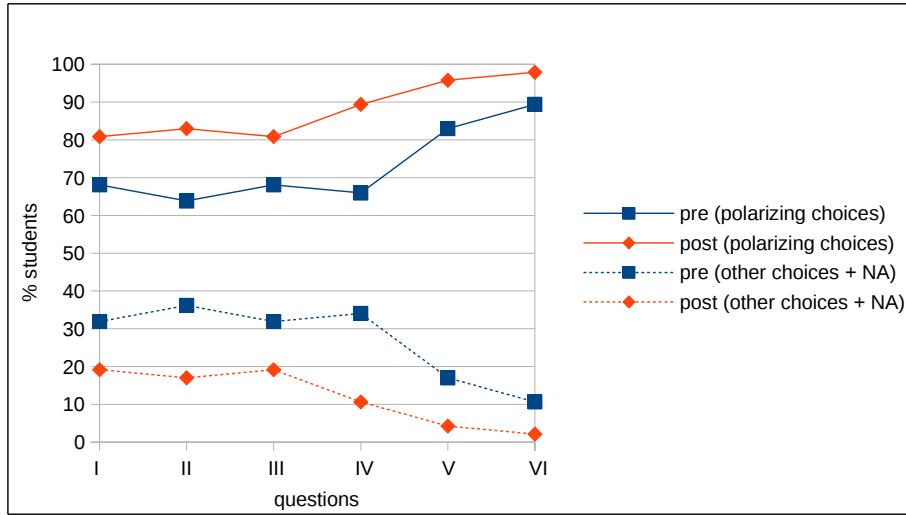
identification of forces acting on a given body. The percent of students who got the correct answer in the posttest is higher than that of the pretest for all questions. However, similar to that of the overall FCI result presented in the previous subsection, the improvement is small. The *normalised gain* expressed in percent ranges only from 13% to 30% which is classified as *low G* in the work of [2] (medium  $G$  :  $30\% < G < 70\%$ , high  $G$  :  $>70\%$ ).



**Figure 2:** Distribution of student answers for questions I to VI. The five choices are labeled A to E (NA means no answer). The letter of the correct answer is labeled with a check mark. The red circles indicate the top two choices “polarising” student answers.

The seemingly uninteresting result above prompted us to look closely on how the students answered the six questions. **Figure 2** shows the result of the item analysis for each question. The number of students who answered each letter for a given question is counted and the result is plotted in the figure. Careful analysis of the results indicate that

- there are two dominant answers for both the pretest and the posttest,
- the pair of dominant answers is the same for both the pretest and the post test for each question,
- the pair of dominant answers contain the correct answer.



**Figure 3:** Distribution of student answers to questions I to VI.

By a “pair of dominant answers,” we mean a pair of choices with the highest frequency. Consider for instance, question I. The pair of dominant answers are choices B and D. For the pretest, the share of the pair of dominant answers is 68% and it rises to 71% in the posttest. Such a large share of the pair of dominant answers justifies our use of the term “dominant”. **Figure 3** shows the wide gap between the percent of students who answered the pair of dominant answers and those who answered the remaining choices for both the pretest and posttest. The percent of students who answered the dominant answers ranges from 64% to 89% for the pretest and from 81% to 98% for the posttest.

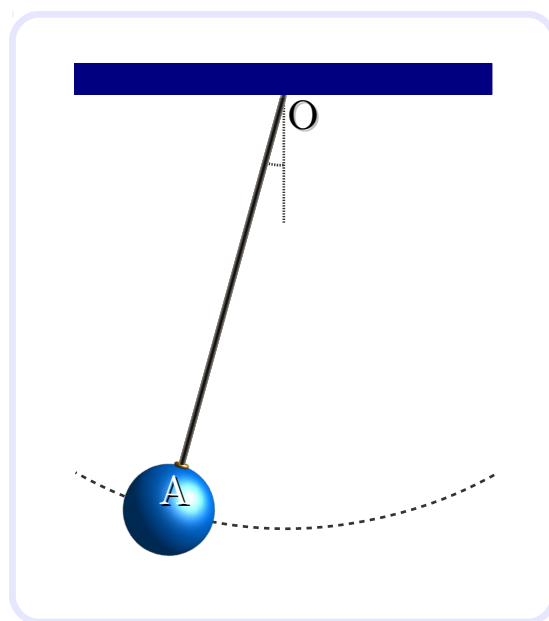
The curious pattern that we found out above regarding the existence of a pair of dominant answers raises the following concerns:

- (a) Why is there a pair of dominant answers?
- (b) What is the nature of this pair of dominant answers?
- (c) Why do students choose the pair of dominant answers?

Observation of the plots in **Fig. 2** in connection to the six FCI questions tells us that the pair of dominant answers always contain the correct choice; symbolically, it is (CC, X) where CC is the correct choice while X is the other choice. Furthermore, both CC and X contains the right set of forces acting on a given body *but* X contains one or more extra forces that do/does not legitimately act on a given body. In other words, X is a superset of CC. The set of extra “fictitious” force(s) in X confuse(s) the student causing a sort of “polarisation”; that is, in analogy to charge polarisation, the majority of student answers are divided into two namely, choices CC and X. For this reason, the more appropriate term for CC and X seems to be *polarising choices* and we are going to use the term from hereon.

The existence of polarising choices brings the idea of misleading ontological categorisation [13]. When facing a problem requiring the identification of forces acting

on a given body, several forces may come to mind. When there are only few forces acting on a given body, it is possible to have a good idea that includes all the valid forces plus the extra forces that are wrong. Whereas the matter of picking up the valid forces and putting it in the category of valid forces may be easily done after a careful analysis, the matter of categorizing the extra forces as correctly acting or not can be quite problematic.



**Figure 4:** A typical question used to test conceptual ability in identifying forces. Consider a pendulum attached at  $O$  swinging from a previous higher point to  $A$ . The students have to distinguish what are “real” and “fictitious” forces from: (i) the downward force of gravity, (ii) the tensional force exerted by the string from  $A$  to  $O$  and (iii) the force in the direction of the pendulum’s motion.

The idea of polarising choices can be made clearer through an example. Consider a snapshot of a swinging pendulum shown in **Fig. 4**. A student may be asked to identify the forces acting on the pendulum bob with the following choices:

- (i) downward force of gravity
  - (ii) the tensional force exerted by the string from  $A$  to  $O$
  - (iii) the force in the direction of the pendulum’s motion
- A. (i) only
  - B. (ii) only
  - C. (iii) only
  - D. (i) and (ii) only
  - E. (i), (ii), and (iii)

The correct answer is letter D corresponding to forces in (i) and (ii). It is worth noting that choice E offers the same set of forces as that of choice D but with an additional

“fictitious” force given by (iii). Choices D and E are our polarising choices for this question. An average student may be able to easily discard choices A, B, and C but may find choices D and E rather confusing. As we have seen in the results above, the inclusion of extra “fictitious” force for questions of this type can cause a serious divide in the answers of students.

One curious thing that caught our attention regarding this matter of polarisation is that most students did well in their homework (average 88%). The immediate question is “How does the system of MasteringPhysics guide the students to the correct answer?” MasteringPhysics [8] has a “corrective measure” by way of hints and trial-and-error. When it comes to answering problems requiring the identification of forces, the existence of hints together with the limited number of tries the students are allowed to make, narrows down the number of possible pathways to the correct answer in their homework. Although we may need more data, our preliminary study here suggests that the “corrective measure” of the system of MasteringPhysics may not be able to correct deeply ingrained student misconceptions/preconceptions about force. However, we are not ignoring the possibility that we might also be able to use such a system to combat the misconception/preconception of extra force(s).

#### 4. Summary and Concluding Remarks

In this article we have discussed several points involving (a) the notion of polarisation as it arises in answering questions about the identification of forces acting on a body, (b) the FCI as an instrument to assess student understanding of forces, (c) the interactive web-based homework and tutorial software MasteringPhysics, and (d) the English language ability as it relates to student learning (see the appendix).

We looked at six particular questions on the FCI to establish if there were any patterns in answering questions. We came across a “polarisation” where student answers tend to cluster about two competing choices with the wrong choice being a superset of the right choice. Our results are suggestive that students have a good idea of the right set of forces acting on a given system but the inclusion of extra force(s) brings about confusion; something that may be explained by misleading ontological categorisation of forces. Progressive hints and limited number of tries offered by interactive web-based Physics education software in answering problems involving proper identification of forces, could help students “get to” the correct answer. This however, may not be enough to significantly alter misleading preconceptions or misconceptions of students.

In the appendix, we raised the issue of English ability and performance on the FCI (more generally any science-based test) and appeared to find *no* correlation, most likely arising from the requirement of a TOEFL score above  $\sim 80$ . In Japan, at least there appears to be no minimum reading or writing score required and the bottom panel of **Figure A1** clearly appears to show a correlation (as might the Total TOEFL score if any range of students were accepted: TOEFL 0  $\sim$  120). This certainly warrants the implementation of a minimum reading and writing score on TOEFL, and similar tests,



for entrance to global courses in Japan.

What we have found out in this study open new avenues for future investigation. Firstly, with regard to the idea of polarising choices, we have so far only scratched the surface of answering intriguing questions such as (a) how students ontologically categorise forces acting on a given body, (b) why polarising choices exist (it may be constructed by design, such as via the use of “distractors” [10], but *only* experiment can tell us whether there really is a resulting polarisation) and how such an idea relates to our way of thinking, (c) how we can “cure” misconceptions/preconceptions related to polarising choices and polarisation, and (d) whether polarising choices and the resulting polarisation can also be induced/found for other physics concepts aside from force. Secondly, the issue of English language ability-FCI for international courses is far from being settled. Although intuitively, there should be a positive-correlation between the two, questions on how far English ability is a factor on learning physics concepts for a multiracial class are something worthy of further investigation.

## Acknowledgements

We would like to thank current and past students on the Chemistry-Biology Combined Major Program (CBCMP) for creating stimulating interactive engagement classes.

## Appendix: English Language vs. Reading Ability

Given that the course is predominantly for non-native speakers of English (the CBCMP entrance requirement is a TOEFL score above  $\sim 80$ ) we considered the possibility that English ability leads to a low  $G$ . We checked to find a very low correlation between pre/post test performance on the FCI and English ability ( $r^2 = 0.0048/0.00041$ , respectively), see **Fig. A1**, including students classed as native.

It may be worth mentioning that the distinction between native and non-native here is blurred because, for administrative reasons a TOEFL/IELTS type score was required for some students who might be considered native (for example even though their language of instruction at high school was English). Actually it is well known that natives do not score full marks on TOEFL/IELTS tests [14]. However, rather than use some kind of complicated probabilistic Monte-Carlo type simulation [15] to randomly model “native” scores for those with TOEFL scores greater than 115 out 120, for simplicity we just give full marks to those classed as “native” administratively. This also applies to reading ability, where in some cases a student is exceptionally better (or worse) than the average taken from their total TOEFL/IELTS score.

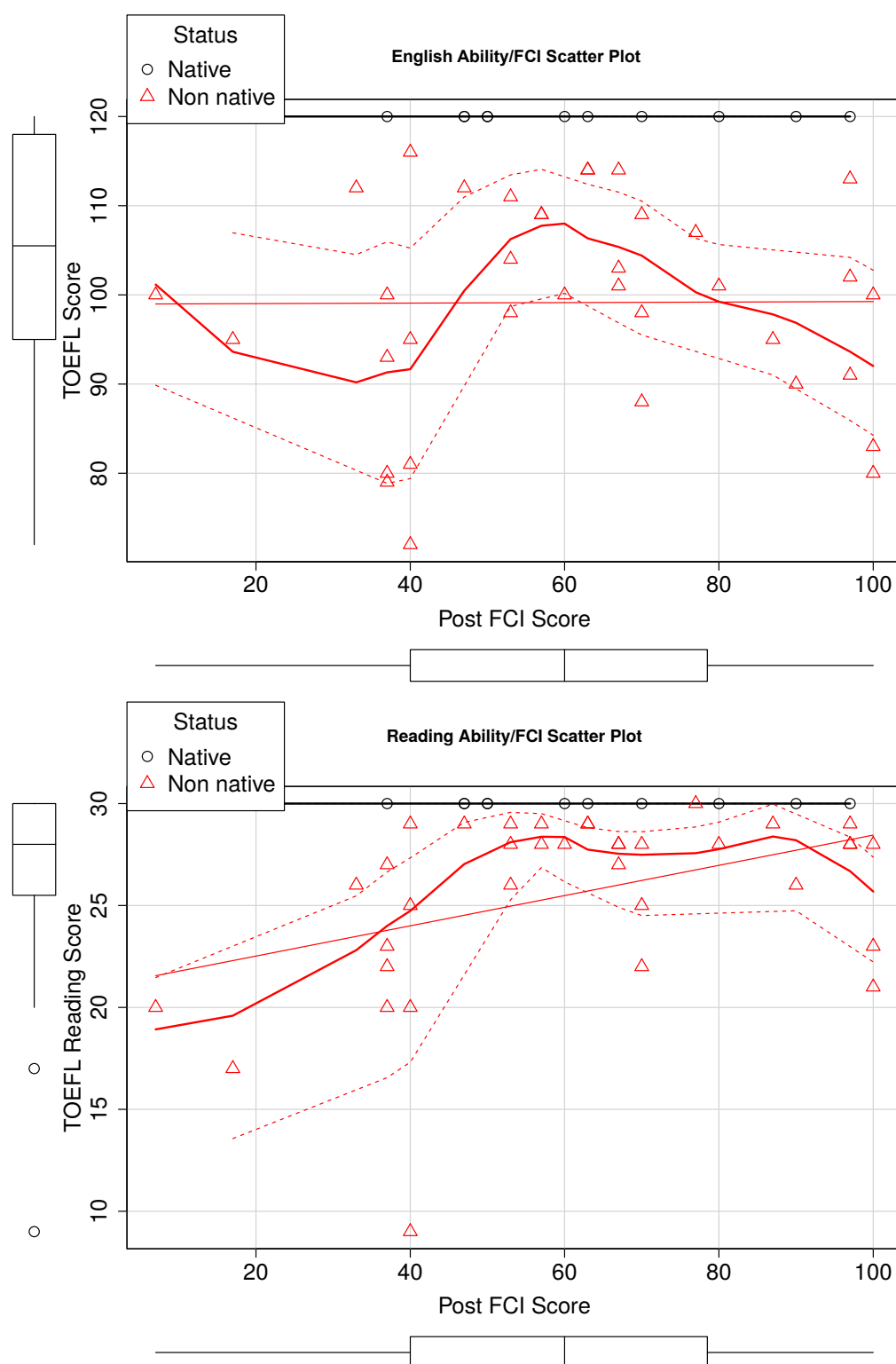
Perhaps not surprisingly we did find a correlation with reading ability on pre/post FCI with  $r^2 = 0.076/0.097$ , respectively. This is an order of magnitude larger than the total TOEFL score correlation. The scatter plot in **Fig. A1** (lower panel) clearly shows some kind of correlation for non-natives in reading ability, with two distinct groups. We stress here that this pattern is not the same problem that occurs for native speakers

in conceptual tests who have confusion between every day and physics usage of words. This can clearly be seen from the fact that we have two separate groups in the scatter plots in **Fig. A1**.

There are then two possible reasons for the fact that the overall score had no correlation. Either, this is because we are taking a small high-end sample of TOEFL in the range 80 – 120, while the entrance requirement did not require a minimum reading score on TOEFL; or that reading ability is the more important factor in conceptual tests on global courses. In Japan most institutions only ask for an overall score on English tests, while the evidence here suggests that lower bounds on ‘reading’ and ‘writing’ ability are very important for students to succeed on university courses in general (we hope to discuss further evidence for this in future work).

## References

- [1] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, *Am. J. Phys.* **60**(7), 627-636, 1991.
- [2] R. R. Hake, Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66** 64-74, 1998.
- [3] A. P. Fagen, C. H. Crouch, and E. Mazur, Peer Instruction : Results from a Range of Classrooms, *Phys. Teach.* **40**, 206-209, 2002.
- [4] E. S. Morote and D. E. Pritchard, What course elements with improvement on tests in introductory Newtonian mechanics, *Am. J. Phys.* **77**, 746, 2009.
- [5] H. Brhane and G. Abebe, Comparison of Teaching Methods in terms of Conceptual Understanding of Newtonian Mechanics: The Case of Freshman Students at Jigjiga University, Ethiopia, *Middle Eastern & African Journal of Educational Research*, **10**, 44-56, 2014
- [6] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, *Phys. Teach.* **30**(3), 141-158, 1992
- [7] A. Savinainen and P. Scott, The Force Concept Inventory : a tool for monitoring student learning, *Phys. Educ.* **37** 45-52, 2002
- [8] MasteringPhysics, Pearson Publishing Group, <http://www.pearsonmylabandmastering.com/masteringphysics/>
- [9] R, R Foundation, <http://www.r-project.org/>
- [10] Adrian Madsen, Sarah B. McKagan, Eleanor C. Sayre, Best Practices for Administering Concept Inventories, arXiv:1404.6500 [physics.ed-ph]
- [11] D. R. Sokoloff and R.K. Thornton, Using interactive lecture demonstrations to create an active learning environment. *Phys. Teach.* **35**(10), 340-347, 1997
- [12] V. P. Coletta and J. A. Phillips, Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability, *Am. J. Phys.* **73**(12), 1172-1182, 2005
- [13] A. T. Johnston and S. A. Southerland, A reconsideration of science misconceptions using ontological categories, *Annual Meeting of the National Association for Research in Science Teaching (New Orleans, Louisiana, April 2000)* on-line at [http://physics.weber.edu/johnston/research/ontological\\_categories\\_reconsidered.pdf](http://physics.weber.edu/johnston/research/ontological_categories_reconsidered.pdf)
- [14] Lawrence J. Stricker and Walter Emmerich, The Performance of Native Speakers of English and ESL Speakers on the Computer-based TOEFL and GRE General Tests, Educational Testing Service, 2002
- [15] C. Robert and G. Casella, Monte Carlo Statistical Methods (2nd ed), Springer, 2005



**Figure A1:** Scatter plot of English (top) & reading ability (bottom) vs posttest FCI. We clearly see two groups: Natives and non natives and although we found no correlation between the overall TOEFL and the post FCI score, there appears to be one for non-natives (similar correlations were found for pretest FCI).